International Journal of Advances in Applied Sciences (IJAAS)
Vol. 11, No. 2, June 2022, pp. 107~112
ISSN: 2252-8814, DOI: 10.11591/ijaas.v11.i2.pp107-112

Improved customer churn prediction model using word order contextualized semantics on customers' social opinion

Ayodeji O. J. Ibitoye[1], Olufade F. W. Onifade[2]
[1]Computer Programme, College of Computing and Communication Studies, Bowen University, Iwo, Nigeria
[2]Department of Computer Science, Faculty of Science, University of Ibadan, Oyo State, Nigeria

Article Info

Article history:

Received Nov 10, 2021
Revised Jan 10, 2022
Accepted Feb 9, 2022

Keywords:

Churn prediction
Customer behavioral analysis
Opinion mining
Semantic analysis
Word order

ABSTRACT

Through the hype in digital marketing and the continuous increase in volume and velocity of opinions about an organization's brands, churn prediction now requires advanced analytics in opinion mining for effective customer behavioral management beyond keywords sentiment analysis (SA). Earlier, by analyzing customers' opinions using SA models, the extracted positive-negative polarity is used to classify customers as churners or non-churner. In those methods, the impact of word order, context, and the inherent semantics of the clustered opinion set were oftentimes overlooked. However, with the consistent creation of new words with new meanings mapped to existing words on the web, the research extended the fuzzy support vector model (FSVM) to show that the dependency distance between the headword, its dependent, and tail word can be weighted by using information content derived from a corpus to generate four-classed social opinion categories as a strongly positive, positive, negative, and strong negative. These opinion classes formed the basis for the churn category as a premium customer, Inertia customer potential churner, and churner in customer behavioral management. In performance evaluation, aside from engendering quadrupled churn class against the existing churn binary pattern, better accuracy, precision, and recall values were obtained when compared with existing SA works in support vector machine and fuzzy support vector machine (FSVM), respectively.

Corresponding Author:

Ayodeji O. J. Ibitoye
Computer Science Programme, College of Computing and Communication Studies, Bowen University
Iwo, Osun State, Nigeria
Email: Ayodeji.ibitoye@bowen.edu.ng

1. INTRODUCTION

In the modern era, the web has become a digital community where different people from different socio-cultural backgrounds and intellectual abilities meet to discuss, argue, criticize, and/or contribute among others to the communities knowledge bank and economic policy [1]. In truth, the digital community initiative has changed serval perspectives through which people reason, act and react especially on social issues. While products and services marketing has gone beyond physical word of mouth to a revolutionized system, the predominant online interaction through social media has now become an effective customer acquisition and retention tool in social network analysis (SNA) [2]. No doubt, the social behavior of customers can influence members of his/her community network. Thus, customer churn remains an instance of user behavior in customer relationship management (CRM) through transactional and/or social network analytics. Goal-driven organizations, especially subscriber-based companies as a necessity try to predict churn to aid proper decision support. Churn, which is used to describe the cessation of a contract, is derived from the words,

change and turn [3]. Among other sectors like bank, game, and employment management, churn prediction remains an important research focus in telecommunication [4]. Churn prediction is more important in telecoms because of the daily growth in the subscription base, competitors' strategies of introducing incentives (promo) to attract and accommodate new customers, and large capital based among others [5]. Research work has indicated that it cost 5–6 times efforts to acquire a new customer than to retain an existing one. Thus, opinion mining, semantic analysis, and other machine learning (ML) models have become critical tools for churn prediction. While the opinion mining method focuses more on identifying positive and negative words expressed in the feelings of a user, semantic analysis deals with the meaning of terms in an expression within a context [6].

However, these methods do not harness the contributions of each word, arrangement, and interdependent relationships that existed between words in a sentence, paragraph, or document for opinion mining. The existing word-independent approaches in opinion mining oftentimes neglect the context of expression, which has become more important in recent times due to the varieties of meanings attached to words and growth in the digital content space. Thus, the goal of the research is to identify and show that the dependent relationship between headwords and their tail equivalents in a sentence can be used to preserve the context of an expression in opinion mining. Here, the impacts of word polarity in a sentence, and its distinctive contributions to the meaning of the sentence are weighted to obtain better sentiment and semantic values of a sentence in opinion mining. Hence, a word order contextual semantics approach is developed to extend the fuzzy support vector machine (FSVM) approach to sentiment analysis (SA). The essence is to effectively analyze customers' opinions on products and brands for a subscriber-based organization towards effective customer behavior management. Therefore, in section 2, a review of existing customer behavioral models through SA is discussed. In section 3, the developed word order contextual semantic approach and algorithm are presented while sample experiments, evaluations, and results of the research are discussed in section 4 before the research was concluded in section 5.

2. CUSTOMER BEHAVIOUR MANAGEMENT THROUGH SENTIMENT ANALYSIS

With increasing competitiveness among organizations in acquiring and retaining customers, the study of customer activities, otherwise known as customer behavior in association with the purchase or services of an individual, or group of people has become an essential factor to aid decision support in CRM. While customer behavior is an inter-disciplinary social science, it also studies how attitudes, emotions, and inclinations determine the behavior of a customer towards a service provider. Many of the customer's attitudes are embedded in expressed opinions, actions, and reactions to the organization's purchase among others. For instance, telecommunication [7], used ML techniques to predict customer behavior for post-paid subscribers since the behavior of a customer toward a service provider, determines the organization's churn value.

Consequently, social websites have provided an unrestricted environment that offers consumers access to product information with a voice to express supporting or counter feelings as such. This, without doubt, can facilitate customers' purchase decisions [8]. But in more recent times, SA has been used in accessing data for product review, movie assessment, trend prediction, medical issues, sports comments, and political issues among others. Examples of popular social media platforms include Blogs, YouTube, MySpace, and Facebook [9]. Based on wider acceptance, availability, and convenience, Twitter, has continued to be a top-rising online social network [10]. It enhances users to express their opinion in short text messages that are not more than 140 characters. Thus, by ML [11] and lexicon-based methodologies [12], SA has been used on tweets for customer behavioral analysis in CRM.

Although, there are three predominantly classification levels for sentiment analysis; document aspect, sentence levels, and [13]. The latter two classes are commonly used for SA in CRM. While sentence-level classification focuses on the voiced sentiment in a sentence, the aspect-level classification goal is to classify sentiments in line with the specific aspects of the entities. By these approaches, SA techniques are applied to different kinds of texts; from blogs [14] to novels [15], newspaper headlines [16], and tweets. In some of the reviewed works on SA, [17] presented a system for objective, positive and negative classification of tweets. Through the corpus, an opinion classifier built on the multinomial Naive Bayes process using features like Ngram and part of speech (POS-tags) was utilized.

However, since it has tweets with only emoticons, the efficiency of the adopted training set was not guaranteed. Also, a fuzzy domain sentiment ontology tree for SA was developed by [18] while a sentiment fuzzy classification algorithm [19] with POS tags was used on a dataset that contains movie reviews to improve the accuracy of the classification. Due to infrequent words available via Twitter, developed the use of microblogging features like emoticons, punctuations, re-tweets, replies, and hashtags, instead of using n-grams [20]. Through this, the classification of the sentiment accuracy was enhanced by 2.2% when equated to

unigrams training via SVM. Furthermore, to perform a distinct classification task, [21] uses dictionaries of negative and positive polarized words, while [12] combined a Naïve Bayes classifier alongside a lexicon-based method to break tweets into negative, neutral positive respectively.

However, many of the ML tactics for detecting the polarity of an opinion still works as domain-specific and temporally dependent [22] while an introduced error via automatic labeling of training data continues to affect the classifiers' performance [23]. Since subjective opinion contains significant bias, a fuzzified support vector approach was used by [24] to ascertain the bias in text and reclassify opinions into dissimilar grades of positive and negative classes by fuzzy membership extension on a SVM. This model accommodated inherent ambiguity in spoken words for better SA results. To generate a more optimal solution especially in customer behavioral management, word order is introduced in this research on FSVM for better SA. Therefore, in section 3, the research subject is discussed.

3. PROPOSED WORD ORDER CONTEXTUAL SEMANTIC FOR OPINION MINING

In determining customers' behavior from expressed opinion, beyond sentiment analysis, the semantic analysis of sentences remains very significant. Thus, to decide the existence of semantic relation between two words, the developed method exploits the inherent connections between words by using dependency parsing to obtain words' semantic relationships. Here, both terms and syntactic words are considered before extracting available sentiments with the influence of word order on various tweets with respect to the context under consideration. Initially, the dependency relations between terms in w_s is extracted before finding the cosine similarity of these terms in w_s. Later, a semantic word order vector score was obtained for each expression in the training set. Therefore, in Algorithm 1, the depending parsing algorithm is presented.

Algorithm 1. Dependency relation

Input: Let t_i be the text that holds plain terms such that $t_i = \{w_0, w_1, w_2, w_3, \ldots \ldots \ldots \ldots . . w_n\}$
Output = Set of concept labels where $C_{lb} = \{c_1, c_2, c_3, \ldots \ldots \ldots \ldots . c_n\}$

1. Let R = $\{r_1, r_2, r_3, \ldots \ldots \ldots . r_n\}$ denote a finite set of functional relations between text
2. Let DG denote the dependency graph where DG = <H, RM> such that
3. H = t_i (set of node to represent the head)
4. RM denotes a set of labelled directed arcs to represent the child where;
5. $RM \subseteq H \times R \times H$
6. For every t_i
7. $< w_i, R, w_j > \in$ RM represents an arc from the head w_i to the dependent w_j labelled with the relation R. Thus
8. if $< w_i, r, w_j > \in$ RM then
9. $< w_x, r', w_j > \notin RM$ for all $x \neq i$
10. elseif $< w_i, r, w_j > \in$ RM then
11. $< w_i, r', w_j > \notin RM$ for all $r' \neq r$
12. else output Subj{w_i to w_n }, pred{w_i to w_n }, obj {w_i to w_n }

Algorithm 2. Churn opinion behavioral analysis

Input: 1 to *n* conversation cluster.
Output: Churn behavioral class and intelligent concept

1. Let D = Lexical Database (here, the sentiWordnet)
2. Let SCT_i) = Current clustered contextual conversation that contains e_i
3. while (type (SCT_i) !=end) do
4. For every indexed tweets e_i in SCT_i
5. Let w_s = word set, where w_s is a set of terms, $t_1 \ldots . . t_n$ in the input variable e_i
6. Identify trigger terms to obtain equivalent senti score t_{sp} in D
7. Apply FMSVM on w_s
8. Extract dependency relations between the terms in w_s
9. Measure the word order contextual semantics of e_i
10. Output e_i sentiment Analysis class
11. Output e_i word order Contextual semantic class
12. Match the outputs in step 10 & 11 to a single equivalence
13. End for
14. End While

While ST_i is the set of all words in C_i [25], and FMSVM is a fuzzy membership support vector model [24]. From Algorithm 1, as different sentences with inherent words gives different meanings, no doubt, to obtain a near accurate semantic, the order of such; unlike with existing semantic analysis system is very important. Thus, to achieve the behavioral analysis of each tweet in a cluster, the entire tweet as an opinion is mined to obtain its sentiment and semantics with the influence of word order. This procedure is presented in Algorithm 2.

4. EXPERIMENTS AND EVALUATION

Through the Twitter streaming application programming interfaces (API), a dataset of 65,315 opinions, with 62,050 tweets having at least an adjective was clustered through keyword search and text streaming on a telecom organization's product, brand, promotions, and other service information. The tweets were pre-processed and redundant data like emotion icons, RT tweets, URLs, and replicated characters were removed. For better analysis, first, the clustered tweets were tokenized and stemmed. Then, the features that define the terms were extracted using the natural language toolkit (NLTK). Consequently, from the obtained clustered dataset, all the POS-labelled words, which corresponded to the seed lexicon are tagged with respect to the lexicon's polarities using SentiWordNet 3.0.

Here, the results of classified opinion with word order influence on fuzzy support vector model are presented in Table 1, later a comparative evaluation between word order fuzzy support vector model (WOFSVM), fuzzy support model, and support vector model is presented in Tables 2 and 3 respectively.

Thus, how many unique customers contributed to each category of opinion classification? In the sequel to equivalence, the following number of customers as presented in Table 5 are obtained for each class of options via a unique customer identity. In all, 75.2% of unique customers are responsible for the 65315-clustered opinions. While 1.9% of opinions are neutral for WOFSVM, 2% and 2.2% are neutral for FSVM and SVM respectively. From the experiments, the WOFSVM performed better than the FSVM and SVM respectively.

Table 1. Opinion classification with WOFSVM

Opinion class	Average percentage cluster	Total number of opinions per class
Strong Positive	9	5879
Positive	14	9144
Neutral	2	1306
Negative	22	14369
Strong Negative	53	34617

Table 2. FSVM opinion classification

Opinion class	Average percentage cluster	Total number of opinions per class
Strong Positive	18	11757
Positive	13	8491
Neutral	2	1306
Negative	21	13716
Strong Negative	46	30045

Table 3. WOFSVM on opinion mining

Approach	Accuracy (%)	Precision (%)	Recall (%)
WOFSVM	84.23	83.17	85.20

Table 4. Percentage performance difference (WOFSVM VS (FSVM and SVM))

Method	Accuracy (%)	Precision (%)	Recall (%)
FSVM	2.49	2.63	2.78
SVM	14.92	11.52	14.53

Table 5. Total number of customers

Churn class	SVM (%)	FSVM (%)	WOFSVM (%)
Premium Customer	8.1	5.6	4.7
Inertia Customer	7.9	7.4	7.2
Potential Churner	24.3	25.6	23.7
Churner	32.7	34.6	37.7

5. CONCLUSION

Beyond SA in the opinion mining is to also pinpoint the semantics of a word in a sentence. Now with the influence of word order, this research has shown that word components can be weighted by using information content derived from a corpus to generate four-classed social opinions class Strong Positive, Positive, Negative, and Strong Negative. This opinion classification was later used to define the new categories of churn i.e. Premium Customer, Inertia Customer. Potential Churner and Churner towards effective decision support for customer behavior management.

REFERENCES

[1] A. K. and T. M. Sebastian, "Sentiment Analysis on Twitter," *International Journal of Computer Science Issues*, vol. 9, no. 3, pp. 1694–16081, 2012.

[2] L. P. Forbes and L. P. Forbes, "Does Social Media Influence Consumer Buying Behavior? An Investigation Of Recommendations And Purchases," *Journal of Business & Economics Research (JBER)*, vol. 11, no. 2, p. 107, 2013, doi: 10.19030/jber.v11i2.7623.

[3] K. Dasgupta *et al.*, "Social ties and their relevance to churn in mobile telecom networks," *Advances in Database Technology - EDBT 2008 - 11th International Conference on Extending Database Technology, Proceedings*, pp. 668–677, 2008, doi: 10.1145/1353343.1353424.

[4] Y. Richter, E. Yom-Tov, and N. Slonim, "Predicting customer churn in mobile networks through analysis of social groups," *Proceedings of the 10th SIAM International Conference on Data Mining, SDM 2010*, pp. 732–741, 2010, doi: 10.1137/1.9781611972801.64.

[5] A. M. Almana, M. S. Aksoy, and R. Alzahrani, "Data Mining Techniques In Customer Churn Analysis For Mobile industry," *Int. Journal of Engineering Research and Applications*, vol. 4, no. 5, pp. 165–171, 2014.

[6] A. Maind, "Measurement of Semantic Similarity Between Words: A Survey," *International Journal of Computer Science, Engineering and Information Technology*, vol. 2, no. 6, pp. 51–60, 2012, doi: 10.5121/ijcseit.2012.2605.

[7] Y. Zhang, R. Liang, Y. Li, Y. Zheng, and M. Berry, "Behavior-based telecommunication churn prediction with neural network approach," *Proceedings - 2011 International Symposium on Computer Science and Society, ISCCS 2011*, pp. 307–310, 2011, doi: 10.1109/ISCCS.2011.89.

[8] R. V. Kozinets, K. De Valck, A. C. Wojnicki, and S. J. S. Wilner, "Networked Narratives: Understanding Word-of-Mouth Marketing in Online Communities," *Journal of Marketing*, vol. 74, no. 2, pp. 71–89, 2010, doi: 10.1509/jm.74.2.71.

[9] S. S. Sin, K. M. Nor, and A. M. Al-Agaga, "Factors Affecting Malaysian young consumers' online purchase intention in social media websites," *Procedia - Social and Behavioral Sciences*, vol. 40, pp. 326–333, 2012, doi: 10.1016/j.sbspro.2012.03.195.

[10] X. L. and T. Z. Long Jiang, Mo Yu, Ming Zhou, "Target dependent Twitter Sentiment Classification," *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 151–160, 2011.

[11] G. Gautam and D. Yadav, "Sentiment analysis of twitter data using machine learning approaches and semantic analysis," *2014 7th International Conference on Contemporary Computing, IC3 2014*, pp. 437–442, 2014, doi: 10.1109/IC3.2014.6897213.

[12] A. P. Pravin Keshav Patil K, "Automatic Sentiment Analysis of Twitter Messages using Lexicon based Approach and Naive Bayes Classifier with Interpretation of Sentiment Variation," *International Journal of Innovative Research in Science Engineering and Technology*, vol. 4, no. 9, pp. 9025–9034, 2015, doi: 10.15680/IJIRSET.2015.0409103.

[13] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014, doi: 10.1016/j.asej.2014.04.011.

[14] A. Neviarouskaya, H. Prendinger, and M. Ishizuka, "Affect Analysis Model: Novel rule-based approach to affect sensing from text," *Natural Language Engineering*, vol. 17, no. 1, pp. 95–135, 2011, doi: 10.1017/S1351324910000239.

[15] T. W. Mohammad, S. M, Yang, "Tracking sentiment in mail: How genders differ on emotional axes," *Proceedings of the ACL Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, WASSA '11*, 2011, doi: 10.48550/arXiv.1309.6347.

[16] J. R. Bellegarda, "Emotion analysis using latent affective folding and embedding," *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, no. June, pp. 1–9, 2010, [Online]. Available: http://www.aclweb.org/anthology-new/W/W10/W10-02.pdf#page=11.

[17] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010*, pp. 1320–1326, 2010, doi: 10.17148/ijarcce.2016.51274.

[18] L. Liu, X. Nie, and H. Wang, "Toward a fuzzy domain sentiment ontology tree for sentiment analysis," *2012 5th International Congress on Image and Signal Processing, CISP 2012*, pp. 1620–1624, 2012, doi: 10.1109/CISP.2012.6469930.

[19] K. Mouthami, K. N. Devi, and V. M. Bhaskaran, "Sentiment analysis and classification based on textual reviews," *2013 International Conference on Information Communication and Embedded Systems, ICICES 2013*, pp. 271–276, 2013, doi: 10.1109/ICICES.2013.6508366.

[20] L. Barbosa and J. Feng, "Robust sentiment detection on twitter from biased and noisy data," *Coling 2010 - 23rd International Conference on Computational Linguistics, Proceedings of the Conference*, vol. 2, pp. 36–44, 2010.

[21] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-basedmethods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, 2011, doi: 10.1162/COLI_a_00049.

[22] J. Read, "Using emoticons to reduce dependency in machine learning techniques for sentiment classification," *ACL-05 - 43rd Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, pp. 43–48, 2005, doi: 10.3115/1628960.1628969.

[23] M. Speriosu, N. Sudan, S. Upadhyay, and J. Baldridge, "Twitter Polarity Classification with Label Propagation over Lexical Links and the Follower Graph," *Workshop on Unsupervised Learning in NLP at the 2011 Conference on Empirical Methods in Natural Language Processing, EMNLP 2011 - Proceedings*, pp. 53–63, 2011.

[24] O. F. O. Ayodeji O.J IBITOYE, "Making Sense out of Social Data: A Generic Fuzzified Support Vector Approach (FSVA) for Public Opinion Classification on Telecoms Data," *IEEE-EECCMC Conference September 8 & 9, 2018: GRT Institute of Engineering and Technology*, 2018.

[25] A. O. Ibitoye, O. F. . Onifade, and C. O. Yinka-Banjo, "Dealing with the paradoxes of customer opinion for effective decision support in churn management," *Challenges and Opportunities for Knowledge Organization in the Digital Age*, pp. 770–785, 2018, doi: 10.5771/9783956504211-770.

BIOGRAPHIES OF AUTHORS

Ayodeji O. J. Ibitoye facilitates at the Department of Computer Science, Bowen University, Iwo, Nigeria. He obtained his B.Sc. and M.sc Computer Science from the prestigious University of Ilorin, Ilorin Kwara State, Nigeria and University of Ibadan, Ibadan, Oyo state, Nigeria in 2009 and 2014 respectively. He completed his Ph.D. in Computer Science, University of Ibadan in January, 2020. He is a young innovative and resourceful Data Scientist with great analytical and exploratory skills. He is a recipient of different local and international fully funded fellowships like the Machine Learning Summer School, London. His research interest is in Big Data Analytics, Information Retrieval, Predictive Technology, Fuzzy Learning, Knowledge Organisation and Decision Support. He is a member of Data Science Nigeria, IAENG, Black in AI among others. He has several peered reviewed publications in journals and conferences in his field of expertise. He serves as the Director, Directorate of Digital Service, Bowen University. He loves singing and writing. He can be contacted at email: ayodeji.ibitoye@bowen.edu.ng.

Olufade F. W. Onifade is Associate Professor of Computer Science, at the University of Ibadan, Oyo State, Nigeria. He obtained a PhD in Computer Science from Nancy 2 University, Nancy, France in 2009 with specialisation in Data Mining, Image Analysis and Pattern Recognition, Predictive Technologies and Information Retrieval. He has published over 90 papers in both local and International referred journals and conferences and has held several fellowships including ETTMIT and the CV Raman Fellowship for African Researchers in India. Dr. Onifade is a member of IEEE, IAENG and CPN. He is the Director of Fadowilly Research Lab. He is the current Deputy Director (Academics), Distance Learning Centre, University of Ibadan, Nigeria. He loves music and travelling. He can be contacted at email: ofw.onifade@ui.edu.ng